

National Rankings as a Means of Evaluating Medical School Library Programs: A Comparative Study

BY NINA W. MATHESON, *Assistant Director**

*Health Information Management Studies
Association of American Medical Colleges
Washington, D.C.*

SUZANNE F. GREFSHEIM, *Chief†*

*Technical Support Services
Himmelfarb Health Sciences Library
George Washington University Medical Center
Washington, D.C.*

ABSTRACT

A comparative study was undertaken to assess the reasons for the low rankings received by George Washington University Medical Center library in the *Annual Statistics for Medical School Libraries in the United States and Canada*. Although internal studies showed the library was successfully satisfying user needs and meeting its primary objectives, the rankings, which include the traditional measures of quality used by accrediting bodies, indicated the contrary. Three hypotheses were postulated to account for the discrepancy. In a matched group of similar libraries: (1) the rankings of an individual library would differ from the national rankings; (2) clustering the variables would change the rankings; and (3) libraries with similar staff size would tend to rank in the same quartile in service and resource variables. All hypotheses were invalidated. Further tests led to the conclusion that the *Annual Statistics* and other traditional measures of quality are inappropriate and inaccurate methods for evaluating library programs, since they only measure resource allocations and not the effectiveness of those allocations. Alternative evaluation methods are suggested.

ONE OF THE obvious temptations following the publication of the *Annual Statistics of Medical School Libraries in the United States and Canada* each year is to judge one's library program by how it ranks in various key categories. Since 1974/75 when the publication of national statistics on medical school library resources and services began, the rankings of George Washington University Medical Center's (GWUMC) library have been rela-

tively static. The library has consistently ranked in the lower third or fourth quartiles on most performance measures.

POSITIVE EVALUATION

Despite this poor showing, when the medical center administrators conducted an evaluation of Educational Support Services among faculty, staff, and students in 1976/77, a "very positive picture" of the library's services, staff, and collection emerged. GWUMC library's primary objectives have always been to serve the educational needs of the students, staff, and faculty, and the information needs of physicians and primary health care staff of George Washington University Hospital and its affiliates. In the study, representatives from all constituent groups within the medical center were asked to evaluate fourteen aspects of the library, including scope and size of collection, interlibrary loan services, circulation policies and services, reference and on-line bibliographic services, and library use instruction. On a scale where "1" represented extreme satisfaction and "5" extreme dissatisfaction, respondents identifying themselves as frequent users gave primarily "1" and "2" ratings to library services and the staff providing them. None of the frequent users registered dissatisfaction and only one "3" was registered in six of the fourteen categories. Respondents registered significant dissatisfaction with photocopy services, the reserve collection, interlibrary loan services, and library use instruction, but in none of these cases did dissatisfied respondents represent more than 22% of the total. The results of the study were interpreted to show that the library had been achieving its objectives and

*Formerly Director, Himmelfarb Health Sciences Library, George Washington University Medical Center, Washington, D.C.

†Reprint requests.

satisfying the information needs of the faculty, staff, students, and primary care personnel within the GWUMC. No evidence to the contrary has surfaced in the years since the evaluation was conducted.

It should be noted that the medical school curriculum relies heavily on library support for self-instructional materials during the first two years of its program. During the final, clinically-oriented years medical students also depend upon the library's book and journal collection for clinical preparation and presentations, research, and case workups. To further test the efficacy of the educational support provided by the library, the match rate for graduates using the National Internship and Residency Matching Program and being accepted into a program which was one of their first three choices was surveyed for 1978-1980 graduates. A high match would indicate a successful educational program and would therefore reflect favorably on the library's program. The match rate between 1978-1980 George Washington University Medical School graduates and their first, second, or third choice internship programs was 81%, 75%, and 76%, respectively. These figures excluded a substantial group of students who were in the military service and who used the military match program. Match rates for this group were also high.

NATIONAL RANKING STUDY UNDERTAKEN

One cannot ignore the national rankings, however, since they, or the raw numbers from which they come, are the criteria used by accrediting bodies in judging a library's program. Library directors themselves have also traditionally used these, or similar numbers, to evaluate existing programs, to develop long-range plans, and to allocate available resources. Therefore, to understand the meaning of the discrepancy between the national rankings and the results of GWUMC's own evaluation and the internship program match rate, a study was undertaken using data from the 1978/79 *Annual Statistics*. The study was designed to test three hypotheses:

- (1) that GWUMC library rankings in a matched group of similar libraries would differ from the national rankings;
- (2) that clustering the variables would alter the GWUMC library's ratings; and
- (3) that libraries with the same number of staff in full time equivalents (FTEs) would tend to rank in the same quartile on the key services and resource variables.

The key categories selected for analysis were the traditional measures used by accreditation bodies to judge the quality of a library's program. Assuming these variables to be justifiable measures of quality, it was expected that at least one of the hypotheses would prove correct and account for the discrepancy. None, however, proved to be valid.

Nine medical school libraries which matched the GWUMC profile were selected from the national group of 128. To be selected, the library had to: (1) be part of a private institution; (2) serve primarily only a school of medicine and allied health; (3) serve the clinical staff of a teaching hospital or clinic; and (4) be located in a major metropolitan area where there was at least one other public or private school of medicine or major health resource library. It was assumed that these characteristics were likely to reflect: (1) a similar mix of basic scientists, clinical researchers, and clinicians; (2) an implied shared resource pool, however informal; and (3) conditions of accessibility common to urban centers.

In addition, data were taken from the *AAMC Curriculum Directory* and the *AAMC Directory of Medical Education* on the number of medical students enrolled, use of self-instruction, and use of part I of the National Board of Medical Examiners (NBME I) examination. It was assumed that there would be a relationship between user group size and some library service measures such as staffing, hours open, and circulation volume. Mean enrollment was 600 and all institutions either used self-instruction or required the NBME I. Most did both. The *NIH Research Grants 1979* directory was the source of data on research activity within the medical schools compared. All schools had active research programs. It was assumed that research grants data would correlate with figures on total volumes, volumes added, library journal subscriptions, interlibrary loan use, and the number of on-line bibliographic searches. The mean dollars in research grant awards for all ten schools was \$5.6 million. GWUMC was either at or near both means (Table 1).

The decision was made to include Johns Hopkins and the University of Pennsylvania despite their higher level of research support than the other eight schools, because many of the factors being tested, especially library services, were not necessarily dependent on income. Admittedly, one would expect factors such as volumes added and serials subscriptions to be affected, but this relationship is not absolute. In fact, a study of the top forty medical schools in terms of research money

TABLE 1
STUDY GROUP

Medical School	State	Students	Research Grants in Millions
George Washington University Medical Center	DC	600	\$ 4.1
Rush Medical College	IL	500	3.5
Tulane University School of Medicine	LA	601	1.5
Johns Hopkins School of Medicine	MD	510	46.7
Tufts University	MA	621	5.3
Mt. Sinai School of Medicine	NY	440	15.7
Jefferson Medical College	PA	903	6.0
Hahnemann Medical College and Hospital	PA	754	3.0
University of Pennsylvania School of Medicine	PA	670	40.0
Vanderbilt University	TN	395	17.0
		Mean 600	Mean 5.6

received in 1978/79 showed that although they were more likely to be institutions with high ranking medical school libraries (57%), ten had libraries in the lower half of the rankings, two of which were in the last quartile. Specifically, although it is frequently assumed that research-oriented institutions necessarily provide larger budgets for library collection development, and therefore have larger collections, a study of the top forty research institutions, in terms of collection size of their libraries, showed that, while 55% did rank in the top third, 12% were in the bottom third and 33% were in the middle. Finally, it was not the attempt of this study to "pick" the competition. The criteria were established and the nine libraries which met these criteria were selected.

The variables analyzed were service factors including circulation figures, interlibrary loans borrowed, interlibrary loans lent, data base searches performed, weekly hours open for use, and reference and directional questions answered; and resource factors including total volumes held, number of volumes added annually, and journal subscriptions.

METHODOLOGY

Ratio and regression analysis, techniques presented in "Standards for university libraries" [1], were adapted to facilitate the comparison of quantitative data. Both analyses use ratios as a basis of comparison. Methodology devised by Cartter for his library resources index, described in the

"Standards," served as a basis and model for the construction of the index tables comparing the selected variables. Cartter's library resources index, an average of three indexes, is computed by first grouping the institutions to be compared and then isolating the three variables: total volumes, volumes added, and journals received. Separate index numbers for each variable are determined by finding the average for each variable and dividing the average value into the value for each institution. The index values for each library are summed, divided by the number of variables in the table, and then sorted in descending order. For example, the average total circulation count for the ten libraries in this study was 42,687. Three libraries had totals of 64,526, 48,869, and 89,739. Dividing the average (42,687) into each of these figures yielded index values of 1.51, 1.14, and 2.10, respectively. Values for the other variables in the service factor cluster were computed for each library, summed, and divided by the number of variables in that cluster to yield an overall service index.

RESULTS

Using this methodology, five tables were constructed. Table 2 clusters Cartter's traditional set of variables used to measure quality: total volume count, net volumes added, and size of the serials subscription list. In this index table, Johns Hopkins ranked highest and Mt. Sinai School of Medicine lowest: GWUMC was seventh.

NATIONAL RANKINGS

TABLE 2

THE LIBRARY RESOURCES INDEX APPLIED TO TEN PRIVATE MEDICAL SCHOOL LIBRARIES, 1978/79

University	Total Volumes Index	Volumes Added Index	Serials Index	Overall Library Resources Index
1. Hopkins	2.44	2.09	1.41	1.98
2. University of Pennsylvania	.81	1.31	1.47	1.20
3. Rush	.87	1.68	1.01	1.19
4. Vanderbilt	1.10	1.49	.94	1.18
5. Thomas Jefferson	1.13	.65	1.14	.97
6. Tulane	1.26	.57	.85	.89
7. George Washington University	.68	.81	.62	.70
8. Hahnemann	.66	.72	.67	.68
9. Tufts	.75	.44	.81	.67
10. Mt. Sinai	.31	.24	1.08	.54

By adding three variables to these three traditional values, a library resources use index (Table 3) was assembled which revealed a slightly different ranking. The circulation and interlibrary loans lent and borrowed values were added to reflect the use of needed resources that were both available and not available. The result was that while Mt. Sinai was last in resources (Table 2), it was fourth in resources use; Jefferson, fifth in resources, was second in resources use; GWUMC dropped to ninth.

Table 4 is a traditional library services index. GWUMC was ranked seventh, mainly on the basis of high data base searches. High-ranking resources and resources use libraries such as Johns Hopkins and the University of Pennsylvania slipped slightly. Mt. Sinai attained its first ranking on the basis of high circulation and high interlibrary loan activity. Interestingly, however, on this scale, some libraries

ranking high in resources and resources use virtually changed places with the smaller libraries.

The tables showed, then, that the first hypothesis was invalid. GWUMC's ranking in a matched group of similar libraries did not differ from its undifferentiated national ranking; it still fell in the third quartile. The second hypothesis was true in only one instance (Table 5), when services were disassociated from measures of collection size and use. In fact, clustering further reduced GWUMC's rankings.

The third hypothesis, that libraries with the same staff size would tend to rank in the same quartile and thus account for GWUMC library's poor showing in the national rankings, was yet to be tested. Table 6 showed that, among the matched group, GWUMC's staff size was consistent with its other rankings, falling within the third quartile. When the resource and services rankings of the five

TABLE 3

LIBRARY RESOURCES USE INDEX

University	Total Volume Index	Volumes Added Index	Serials Index	Interlibrary Loans Borrowed	Interlibrary Loans Lent	Circulation Index	Overall Index
1. Hopkins	2.44	2.09	1.41	.86	.73	1.51	1.51
2. Thomas Jefferson	1.13	.65	1.14	1.03	2.17	2.03	1.36
3. Rush	.87	1.68	1.01	1.74	—	1.14	1.29
4. Mt. Sinai	.31	.24	1.08	1.61	1.63	2.10	1.16
5. Tulane	1.26	.57	.85	.66	2.01	1.06	1.07
6. University of Pennsylvania	.81	1.31	1.47	1.19	.25	.47	.92
7. Vanderbilt	1.10	1.49	.94	.56	.48	.56	.86
8. Tufts	.75	.44	.81	.87	.74	.23	.64
9. George Washington University	.68	.81	.62	.74	.54	.39	.63
10. Hahnemann	.66	.72	.67	.73	.45	.50	.62

TABLE 4
LIBRARY SERVICES INDEX 1978/79

University	Interlibrary Loans Borrowed Index	Interlibrary Loans Lent Index	Circulation Index	Data Base Searches Index	Weekly Hours Open Index	Overall Service Index
1. Mt. Sinai	1.61	1.63	2.10	.56	1.08	1.40
2. Thomas Jefferson	1.03	2.17	2.03	.61	1.05	1.38
3. Tulane	.66	2.01	1.06	1.45	.95	1.23
4. Rush	1.74	—	1.14	.74	.95	1.14
5. Hopkins	.86	.73	1.51	1.42	.86	1.08
6. University of Pennsylvania	1.19	.25	.47	2.20	1.02	1.03
7. George Washington University	.74	.54	.39	1.51	1.11	.86
8. Vanderbilt	.56	.48	.56	.64	1.04	.66
9. Tufts	.87	.74	.23	.38	.97	.64
10. Hahnemann	.73	.45	.50	.50	.96	.63

TABLE 5
LIBRARY SERVICES AND ACCESS INDEX

University	Data Base Searches	Reference Transactions	Weekly Hours Open	Overall Index
1. Tufts	.38	2.97	.97	1.44
2. Thomas Jefferson	.61	2.31	1.05	1.32
3. Hopkins	1.42	1.28	.86	1.19
4. University of Pennsylvania	2.20	.30	1.02	1.17
5. George Washington University	1.51	.87	1.11	1.16
6. Vanderbilt	.64	1.29	1.04	.99
7. Tulane	1.45	.05	.95	.82
8. Mt. Sinai	.56	.69	1.08	.78
9. Rush	.74	.06	.95	.58
10. Hahnemann	.50	.19	.96	.55

TABLE 6
STAFF SIZE RANKINGS OF TEN PRIVATE MEDICAL SCHOOLS, 1978/79

University	Professional	Nonprofessional	Total Staff (FTE)
1. Hopkins	10	26	36.0
2. Mt. Sinai	11	22.4	33.4
3. Thomas Jefferson	10	22.4	32.4
4. Vanderbilt	8.8	22.4	31.0
5. Tufts	5	23.4	28.4
6. George Washington University	8	19.2	27.2
7. University of Pennsylvania	9	14	23.0
8. Rush	5	17	22.0
9. Hahnemann	5	15	20.0
10. Tulane	6	9	15.0

NATIONAL RANKINGS

TABLE 7

RANK ORDER SCORES FOR LIBRARIES WITH EQUIVALENT FTE STAFF (N = 128)

George Washington University Quartile Rank		Indiana University	Albany Medical College	Stanford University	George Washington University Medical Center	SUNY Stony Brook	University of California, Irvine
3	FTE Staff	(28)	(27.7)	(27.6)	(27.24)	(26.8)	(26.48)
4	Volumes in collection	59	95	10	102	36	66
4	Net volumes added	86	88	92	112	8	60
4	Journal subscriptions	66	100	30	109	25	32
2	Interlibrary loans borrowed	63	9	105	40	107	122
4	Interlibrary loans lent	19	62	35	101	52	59
1	Hours open	82	47	14	11	12	39
3	Hours staffed by professionals	86	41	67	83	31	99
4	Circulation	45	95	40	107	90	112
2	Reference transactions	6	29	6	56	59	50
2	Data base searches	8	36	67	38	45	90

libraries whose staff fall within the same range as GWUMC were compared, however, no pattern was evident. Table 7 shows a wide variation of rankings among variables. Aside from proving the third hypothesis invalid, all that could be concluded was that among libraries reporting similar staff size, GWUMC had the smallest collection, added the least to the collection, loaned the least, had nearly the lowest circulation, and recorded a comparatively low number of reference transactions.

CONCLUSION

The results lead to one of two possible conclusions: either quantitative variables are inappropriate measures of quality when unrelated to a library's goals and objectives, or despite its users' contentions, GWUMC's library is not a very responsive, useful medical school library. The first conclusion can be reached by hypothesizing that the number of volumes variable may reflect a library's weeding policy, not a varied and vital current collection. Also, varying circulation figures can be a function of varying circulation policies and not necessarily collection utility. The editorial board of the *Annual Statistics* acknowledged in the introduction to the 1978/79 volume that there is a "lack of uniformity in data collection among medical libraries, and evidence of an inability to agree upon definitions of standard indicators" [2], which calls into question the usefulness of these

statistics in making valid comparisons. In addition, a similar study conducted by the National Center for Education Statistics (NCES) on sixty-two public and private academic libraries to determine whether qualitative evaluation was possible using the quantitative data provided by the Library General Information Survey concluded that ratio analyses based on quantitative measures were not helpful [3].

Because of the questionable nature of the variable rankings as a means of evaluating a medical school library's programs, a Spearman rank correlation test was conducted which matched library rankings in the overall library resources index with the number of medical students enrolled in each of the schools. No correlation was found. The overall library resources use rankings were then matched against the amount of money in federal funding each school received. Correlation significant at the .05 level existed between the rankings and money for research. One might conclude that the library's rankings have meaning only if the medical school has a strong emphasis on research. But when a library's objectives emphasize education and clinical or patient care priorities, the rankings do not provide a satisfactory means of measuring effectiveness. It must be pointed out, however, that variables analyzed in this study are the yardsticks used by accreditation committees when judging a medical school library's education support capabilities.

From our own study and the NCES study, it is evident that all we are able to do now with the statistics is merely measure resource allocations. What is needed is a means to evaluate the effectiveness of that allocation by determining its worth and its impact on users. Alternative approaches or other statistics which can be used for evaluation must be found. The NCES study rejects the notion of formula standards as an alternative since no justification for base figures is provided [3] and again no qualitative judgments about the choices or decisions represented by the quantitative data are possible. An assessment of what libraries do, or can do, rather than what they have, or should have, is necessary.

Figures collected should not only assess the increments of growth and activity changes, but should provide data on an academic health sciences library's capability, performance, and planning in the areas of service, technology support systems, personnel support, and research and development. The NCES study also suggests that "user satisfaction [studies] based on the performance of the acquisition policy, circulation policy, library operations (e.g. that books are shelved promptly and in the proper location), and user performance are more effective, informative measures for purposes of evaluation." As an added benefit, such studies can provide not only data for comparative purposes, but a means of discovering where performance needs to be improved [3].

Until either more meaningful, standardized data

are collected or some correlation between the statistics and library objectives can be ascertained, it is our opinion that a valid analysis of a library's program can only come from a study of internal criteria. Analyses must therefore be assessments of the validity of the goals and objectives the library has established within its institutional and environmental framework, and the progress made toward the achievement of these goals and objectives. They cannot be based merely on undifferentiated numbers which do not consider variations in policies and record keeping, or user satisfaction and expectation. Possibly, future studies reported in the literature will prove us wrong and provide evidence that libraries are able to use the statistics and rankings for meaningful comparative purposes. It is more likely, however, that different quantitative measures will have to be developed before meaningful comparisons can be made.

REFERENCES

1. Standards for university libraries. CRL News 1979 Apr; 4:101-10.
2. Annual statistics of medical school libraries in the United States and Canada 1978/79. Houston TX: Houston Academy of Medicine—Texas Medical Center Library, 1980:i.
3. Bentley S. Academic library statistics: a search for a meaningful evaluative tool. Libr Res 1979; 1:143-52.

Received March 17, 1980; revision accepted January 7, 1981.